

Retrospective Review
Twitter, Inc. and the 2018 Midterm Elections in the United States

January 31, 2019
Updated: February 4, 2019



Table of Contents

I. BACKGROUND	2
II. LESSONS LEARNED FROM THE 2016 ELECTION	3
A. Retrospective Review	3
B. Election Integrity Data Disclosure	4
C. Insights from Our Review	5
III. IMPROVEMENTS TO TWITTER FOR THE 2018 ELECTION	7
A. Combating Malicious Automation and Protecting Conversation Health	7
B. Cross-Functional Team	8
C. Advertising and Promoted Content	9
D. Recent Updates to Twitter Rules	11
E. Additional Safety Measures for Accessing Twitter's Application Programming Interfaces	12
IV. ENGAGEMENT WITH KEY STAKEHOLDERS	14
A. Government Entities	14
B. Industry Peers	14
C. Candidate Verification	15
D. Election Labels	15
E. Civil Society	16
V. ACTIVITY ON THE SERVICE	17
A. Voter Registration Efforts	17
B. News Broadcasts	19
C. Monitoring and Identification of Violative Tweets	19
1. External Identification of Problematic Content	19
2. Internal Identification of Problematic Content	20
3. Examples of Violative Content	20
D. Malicious Accounts Located in Russia	22
E. Malicious Accounts Located in Iran	22
F. Malicious Accounts Located in Venezuela	24
VI. CONCLUSION	25



I. BACKGROUND

The purpose of Twitter is to serve the public conversation. We serve our global audience by focusing on the needs of the people who use our service, and we put them first in every step we take. People treat us like a global town square, where people from around the world come together in an open and free exchange of ideas. We must be a trusted and healthy place that supports free and open democratic debate.

Twitter is committed to improving the collective health, openness, and civility of public conversation on our service. Twitter's health is built and measured by how we help encourage more healthy debate, conversations, and critical thinking. Conversely, abuse, malicious automation, and manipulation detract from it. In September 2018, Chief Executive Officer Jack Dorsey reiterated Twitter's commitment to be held publicly accountable toward progress, and this retrospective review is part of our ongoing efforts to provide greater transparency to the American people and to individuals around the globe.

The public conversation occurring on Twitter is never more important than during elections, the cornerstone of democracies across the globe. Our service shows the world what is happening, democratizes access to information and—at its best—provides people insights into a diversity of perspectives on critical issues; all in real time. We work with commitment and passion to do right by the people who use Twitter and the broader public. Any attempts to undermine the integrity of our service are antithetical to our fundamental rights and erode the core tenets of freedom of expression, the value upon which our company is based. This issue affects all of us and is one that we care deeply about as individuals, both inside and outside the company.



II. LESSONS LEARNED FROM THE 2016 ELECTION

Twitter continues to engage in intensive efforts to identify and combat state-sponsored and non-state sponsored hostile attempts to abuse social media for manipulative and divisive purposes. We now possess a deeper understanding of both the scope and tactics used by malicious actors to manipulate our service and sow division across Twitter more broadly. Our efforts enable Twitter to fight this threat while maintaining the integrity of peoples' experience on the service and supporting the health of conversation on our service. Our work on this issue is not done, nor will it ever be. It is clear that information operations and coordinated inauthentic behavior will not cease. These types of tactics have been around for far longer than Twitter has existed — they will adapt and change as the geopolitical terrain evolves worldwide and as new technologies emerge. As such, the threat we face requires extensive partnership and collaboration with government entities, civil society experts and industry peers. We each possess information the other does not have, and our combined effort is more powerful in combating these threats together.

A. Retrospective Review

In fall 2017, we conducted a comprehensive retrospective review of potential service manipulation activity related to the 2016 election. This analysis was divided into two parts: (1) a review of organic activity that included investigations into both the Russian Internet Research Agency specifically and broader malicious automation originating in Russia; and (2) a comprehensive review of promoted election-related Tweets linked to Russia.

First, to better understand the nature of the threat of malicious automation and identify ways to address future attempts at manipulation, we examined activity on the service during the 2016 election period. We focused on identifying accounts that were automated, potentially linked to Russia, trying to get unearned attention, and Tweeting election-related content, comparing activity by those accounts to overall activity on the service during the election as a baseline.



As we reported in January 2018, we identified 50,258 automated accounts that were Russian-linked and Tweeting election-related content, representing less than two one-hundredths of a percent (0.016%) of the total accounts on Twitter at the time. Of all election-related Tweets that occurred on Twitter during that period, these malicious accounts constituted approximately one percent (1.00%), totaling 2.12 million Tweets. Additionally, in the aggregate, automated, Russian-linked, election-related Tweets from these malicious accounts generated significantly fewer impressions (i.e., views by others on Twitter) relative to their volume on the service.

Twitter is committed to ensuring that promoted accounts and paid advertisements are free from bad faith actors, including foreign state actors seeking to manipulate our service. In connection with the work we did in the fall of 2017, we conducted a comprehensive analysis of accounts that promoted election-related Tweets on the service throughout 2016 in the form of paid ads. We reviewed nearly 6,500 accounts and our findings showed that approximately one-tenth of one-percent—only nine accounts—were Tweeting election-related content and linked to Russia. The two most active accounts out of those nine were affiliated with Russia Today (“RT”), which Twitter subsequently barred from advertising on Twitter. And Twitter is donating the \$1.9 million that RT spent globally on advertising to academic research into election and initiatives related to elections and civic engagement. The first recipients of those funds include the Kofi Annan Foundation’s Global Commission on Elections, Democracy, and Security, the Atlantic Council, the EU DisinfoLab and the Reporters Committee for Press Freedom.

B. Election Integrity Data Disclosure

On October 17, 2018, Twitter publicly released all the accounts and related content associated with potential information operations that we had found on our service, including the accounts we believe were associated with the activities of the Internet Research Agency on Twitter dating back to 2009. Twitter had previously disclosed these activities, but through these efforts we released substantially more information about them to enable independent research and investigation.



Prior to the release of these datasets, Twitter shared examples of alleged foreign interference in political conversations on Twitter by the Internet Research Agency (IRA) and provided the public with a direct notice if they interacted with these accounts. In August, we also disclosed details of another attempted influence campaign we identified as potentially located within Iran. We provided details about these accounts to Congressional committees and law enforcement, as well as to our peer companies.

In line with our strong principles of transparency and with the goal of improving understanding of foreign influence and information campaigns, we released the full, comprehensive archives of Tweets and media connected with two previously disclosed and potentially state-backed operations on our service. We made this data available with the goal of encouraging open research and investigation of these behaviors from researchers and academics around the world.

These large datasets comprise 3,613 accounts affiliated with the IRA, originating in Russia, and 770 other accounts, potentially originating in Iran. They include more than 10 million Tweets and more than 2 million images, GIFs, videos, and Periscope broadcasts, including the earliest on-Twitter activity from accounts connected with these campaigns, dating back to 2009.

Twitter believes that independent analysis of this activity by researchers is a key step toward promoting shared understanding of these threats. We strongly believe that this level of transparency can enhance the health of the public conversation on the internet.

C. Insights from Our Review

The process of investigating suspected foreign influence and information campaigns is an ongoing one. Although the volume of malicious election-related activity that we could link to Russia was relatively small, we strongly believe that any such activity on Twitter is unacceptable. We remain vigilant about identifying and eliminating abuse on the service perpetrated by hostile foreign actors, and we will continue to invest in resources and leverage our technological capabilities to do so.



Twitter continues to identify service manipulation campaigns focused on political themes, bearing behavioral similarities to other known information operations. Twitter recently has removed 764 accounts originating in Venezuela, some of which Tweeted about the 2016 U.S. election. In total, these 764 accounts Tweeted 984,980 times combined, most commonly in English. On average, each account had 591 followers and Tweeted approximately 1,300 times. The engagement rate for Tweets issued from these accounts were low with only 0.3 “likes” and 0.2 retweets on average. There were no promoted Tweets, and no money spent by these accounts on advertising. Of the 764 accounts, 588 were suspended prior to November 2017. As we have previously noted, due to the use of proxy servers, virtual private networks (VPNs), and other identity-masking technologies, attribution of activity on Twitter to an external actor is challenging, and we rely on our partnerships with government agencies and law enforcement to improve our understanding of these issues. At this time, Twitter is unable to tie the accounts originating in Venezuela directly to a foreign government. Following the public release of additional information operations datasets on January 31, 2019, an external researcher provided information that informed us that we initially misidentified 228 accounts as connected to Russia. As our investigations into the account activity continued, we uncovered additional information allowing us to more confidently associate them with Venezuela. This report has been updated to reflect the most accurate attribution information we currently possess.

We also recognize that, as a private company, there are threats that we cannot understand and address alone. We must continue to work together with elected officials, government entities, industry peers, outside experts, and other stakeholders so that the American people and the global community can understand the full context in which these threats arise.



III. IMPROVEMENTS TO TWITTER FOR THE 2018 ELECTION

We have made the health of Twitter our top priority and our efforts will be measured by how we help encourage more healthy conversations on the service. Conversely, abuse, automation, and manipulation detract from the health of our service. Since January 2017, we have launched approximately 70 product changes and experiments and dozens of new policy changes, expanded our enforcement and operations, and strengthened our team structure, all designed to foster the health of the service and protect the people who use Twitter from abuse and malicious automation. Twitter has made a number of improvements specifically in preparation the 2018 election, described below.

A. Combating Malicious Automation and Protecting Conversation Health

Using the insights from our retrospective review, Twitter continues to develop the detection tools and systems needed to combat malicious automation on our service. For example, Twitter prioritizes identifying suspicious account activity, such as exceptionally high-volume Tweeting with the same hashtag or mentioning the same @handle without a reply from the account being addressed; where we identify such activity, we may require an individual using the service to confirm human control of the account or their identity identity. Twitter has also increased its use of challenges intended to catch automated accounts, such as reCAPTCHAs, that require individuals to identify portions of an image or type in words displayed on screen, and password reset requests that protect potentially compromised accounts. Twitter has also implemented mandatory email or phone verification for all new accounts.

Our efforts have become increasingly effective. Due to technology and process improvements, we are now removing 214% more accounts year-over-year for violating our service manipulation policies. For example, over the course of the last several months, our systems identified and challenged between 8.5 million and 10 million accounts each week suspected of misusing automation or producing spam. This includes many forms of automated account interactions and behaviors as well as attempts to mislead or deceive people. This constitutes more than three times the 3.2 million accounts we challenged per week in September 2017. Additionally, we thwart 530,000 suspicious logins a day, approximately double the amount of potentially malicious logins that we detected and prevented a year ago.



Our motivation with these changes is to reduce the burden on people on Twitter to report spam and malicious automation. These technological improvements have brought about a corresponding reduction in the number of spam reports from people on Twitter. We received approximately 25,000 such reports per day in March of this year; that number decreased to 17,000 per day in August. First published on July 2, 2012, our biannual Twitter Transparency Report highlights trends in legal requests, intellectual property-related requests, and terms of service enforcements. We recently revised the Transparency Report to include findings regarding automated manipulation, and we will continue to share updates on an ongoing basis about our efforts to combat malicious automated manipulation.

We also removed locked accounts from follower counts to ensure these figures more accurately reflect the actual reach and popularity of an account. Accounts are locked when our systems detect unusual activity and force a password change or other challenge. If the challenge has not been met or the password has not been changed within a month, the account is locked, barring it from sending Tweets, Retweets or liking posts from others. Removing locked accounts from follower counts helps ensure that people can trust the reliability of information on profiles on Twitter.

B. Cross-Functional Team

Our improvements in preparation for the 2018 U.S. midterm elections included important structural changes. Twitter Chief Executive Officer Jack Dorsey recently reorganized the structure of the company to allow employees greater durability, agility, invention, and entrepreneurial drive. The reorganization simplified the way we work, and enabled all of us to focus on the health of our service.

In particular, Twitter created an internal cross-functional analytical team whose mission is to monitor site and service integrity. Drawing on expertise across the company, this team can respond immediately to escalations of inauthentic, malicious automated or human-coordinated activity on the service. The team's work enables us to better understand the nature of the malicious activity and mitigate it more quickly.

To supplement its own analyses, Twitter's analytical team also receives and responds to reports from across the company and from external third parties. The results from all of the team's analyses are shared with key stakeholders at Twitter and provide the basis for policy changes, product initiatives, and the removal of accounts.



The primary focus of the cross-functional analytical team is election readiness. Leading up to and during the 2018 election period, the team examined, responded to, and escalated instances of suspected inauthentic, election-related coordinated activity in political conversation and conduct in-depth analyses of relevant Twitter data.

Among the development of numerous additional tools, our cross-functional team developed a political conversations dashboard to surface information about sudden shifts in sentiment around a specific conversation, suggesting a potential coordinated campaign of activity, as well as information about groups of potentially linked accounts that are posting about the same topic.

Through real-time review and detection of anomalous and potentially malicious automated and human-coordinated activity, the team worked to identify and address any attempts by bad faith actors to interfere with the electoral process, and was better informed about where and how to deploy resources to proactively review potential malicious activity. Accounts were escalated for review in real time if exhibiting anomalous patterns of behavior. These efforts significantly improved our ability to detect malicious automated and human-coordinated activity surrounding political content as well as the speed with which we address those issues. This team also responded to reports of malicious activity flagged by third-party partners, as discussed in section IV.

C. Advertising and Promoted Content

As we learned from our 2016 retrospective review, bad faith actors have attempted to influence the electoral process by propagating paid content on the service, including political advertisements and promoted Tweets. As we reported in the fall of 2017, we have devoted considerable resources to increasing transparency and promoting accountability in the ads served to Twitter customers.

Twitter implemented an updated Political Campaigning Policy in June 2018 to provide clearer guidance about how we define political content and who can promote political content on our service. Under the revised policy, advertisers who wish to target the United States with federal political campaigning advertisements are required to self-identify as such and certify that they are located within the United States. Foreign nationals are not be permitted to serve political ads to individuals who identify as being located in the United States.



Twitter account holders who wish to target the U.S. with federal political campaigning advertisements must also comply with a strict set of requirements. Among other things, the account's profile photo, header photo, and website must be identical to the individual's or organization's online presence. In addition, the advertiser must take steps to verify that the address used to serve advertisements with content related to a federal political campaign is genuine.

To further increase transparency and better educate those who access promoted content, accounts serving promoted Tweets with content related to a federal political campaign is now visually identified and contains a disclaimer. This feature allows people to more easily identify federal political campaign advertisements, quickly identify the identity of the account funding the advertisement, and immediately tell whether it was authorized by the candidate.

In June, we launched the Ads Transparency Center, which is open to everyone on Twitter and the general public, and which includes comprehensive data on paid electioneering communications on the service within the United States. Twitter requires extensive information disclosures of any account involved in federal electioneering communications and provides specific information to the public via the Ads Transparency Center, including:

- Purchases made by a specific account;
- All past and current ads served on the service for a specific account;
- Targeting criteria and results for each advertisement;
- Number of views each advertisement received; and
- Certain billing information associated with the account.

These are meaningful steps that have enhanced the Twitter experience and protected the health of political conversations on the service.



We also implemented the next phase of our efforts to provide transparency with the launch of a U.S.-specific Issue Ads Policy and certification process in September 2018. In addition to the policy governing advertisers running campaigns for federal elections described above, Twitter implemented a new disclosure policy for advertisers promoting content about candidates running for federal, state, or local election, as well as those discussing issues of legislative national importance. To provide people with additional information about individuals or organizations promoting issue ads, Twitter has established a process that verifies an advertiser's identity and location within the United States.

These advertisements will also be included in the Ads Transparency Center. We are also examining how to adopt political campaigning and issue ads policies around the world. We remain committed to continuing to improve and invest resources in this space.

As of December 3, 2018, Twitter received 407 applications to register as political advertisers and 544 applications to register as issue advertisers in the United States. In total, we fully approved 236 political advertisers and 332 issue advertisers. We also received 28 applications deemed to be a news outlet thereby exempting the organizations from the Ads Transparency Center requirements.

In 2018, 96 political advertisers spent nearly \$2.3 million to purchase 2,267 ads that resulted in nearly 170 million impressions. Additionally, 150 issue advertisers spent \$2.2 million on 1,373 ads that resulted in approximately 198 million impressions.

D. Recent Updates to Twitter Rules

On October 1, 2018, Twitter announced an update the Twitter Rules to provide clearer guidance around several key issues impacting the integrity of elections across the globe. As service manipulation tactics continue to evolve, we are updating and expanding our rules to better reflect how we identify fake accounts, and what types of inauthentic activity violate our guidelines. Some of the factors that we take into account when determining whether an account is fake include the use of stock or stolen avatar photos; the use of stolen or copied profile bios; and the use of intentionally misleading profile information, including profile location.



We also updated the Twitter Rules regarding attributed activity. Now, if we are able to reliably attribute an account on Twitter to an entity known to violate the Twitter Rules, we will remove additional accounts associated with that entity. We are expanding our enforcement approach to include accounts that deliberately mimic or are intended to replace accounts we have previously suspended for violating our rules. These steps allow us to take more aggressive action against known malicious actors, such as the Russian Internet Research Agency.

Additionally, we announced an update regarding hacked materials. Twitter rules prohibit the distribution of hacked material that contains private information or trade secrets, or could put people in harm's way. According to the Twitter Rules, Twitter does not permit the use of our services to directly distribute content obtained through hacking that contains personally identifiable information, may put people in imminent harm or danger, or contains trade secrets. Direct distribution of hacked materials includes posting hacked content on Twitter (for instance, in the text of a Tweet, or in an image), or directly linking to hacked content hosted on other websites.

We also expanded the criteria for when we will take action on accounts which claim responsibility for a hack, which includes threats and public incentives to hack specific people and accounts. We also may suspend accounts in which Twitter is able to reliably attribute a hack to the account distributing that content. Commentary about a hack or hacked materials, such as news articles discussing a hack, are generally not considered a violation of this policy. This includes, for example, journalistic and editorial discussion of a hacking and disclosures of legitimate public concern and which pose no physical harm.

E. Additional Safety Measures for Accessing Twitter's Application Programming Interfaces

To further address malicious automation and abuse on the service, we have also recently updated our developer policies and processes, which govern the access and use of public Tweet data made available to developers and other third parties through our application programming interfaces ("APIs").



We recognize that programmatic access to the Twitter service, including access to public Tweet data, could be manipulated, so we have taken steps to prevent the use of our APIs for products and services that are abusive or that disrupt the health of conversations. Applications to which we grant access to our APIs are prohibited from using the data to manipulate conversations or otherwise disrupt the integrity of the Twitter service or invade the privacy of people on Twitter. Between July and December 2018 alone we removed more than 162,000 applications that we determined to be in violation of our developer policies. Most violated our policies against producing spam via APIs. And we continue to invest in and improve our detection tools to stop misuse of public Twitter data.

In July 2018, we introduced a new measure designed to increase developers' accountability for applications that create and engage with Twitter content and accounts. Twitter now reviews and conducts compliance checks of all developers' stated use of the data that they wish to access. We have also added new protections aimed to prevent the registration of low quality and spam-generating applications. We believe that these additional steps will help protect the integrity of our service.



IV. ENGAGEMENT WITH KEY STAKEHOLDERS

Information sharing and collaboration are critical to Twitter's success in preventing hostile foreign actors from disrupting meaningful political conversations on the service.

A. Government Entities

We have well-established relationships with law enforcement agencies active in this arena, including the Federal Bureau of Investigation Foreign Influence Task Force and the U.S. Department of Homeland Security's Election Security Task Force. We look forward to continued cooperation with federal, state, and local government agencies on election integrity issues, because in certain circumstance only they have access to information critical to our joint efforts to stop bad faith actors.

On Election Day, Twitter virtually participated in an operations center convened by the U.S. Department of Homeland Security. The operations center also convened officials from the U.S. Department of Justice, the Federal Bureau of Investigation, and the Office of the Director of National Intelligence, in addition to federal, state, local, and private sector partners. In the lead up to Election Day, and throughout the course of the day itself, Twitter remained in constant contact with officials throughout all levels of government.

We also worked in close collaboration with the National Association of Secretaries of State (NASS) and the National Association of State Election Directors (NASED). Founded in 1904, NASS is the nation's oldest, nonpartisan professional organization for public officials, and is open to secretaries of states and lieutenant governors in the 50 states, D.C. and territories. In February, Twitter participated in a panel discussion convened by NASS on the Role of Social Media in Democracy and their New Voters Forum, broadcast on C-Span.

B. Industry Peers

We recognize the value of inputs we receive from our industry peers about hostile foreign actors. We have shared and remain committed to sharing information across services to better understand and address the threat of hostile foreign interference with the electoral process. On August 24, 2018, Twitter hosted our industry peers to discuss data sharing about hostile foreign actors regarding 2018 election security.



Twitter remained in close contact with our industry peers in the lead up to Election Day and throughout the day itself. Further, a number of technology companies — including Twitter — established a dedicated, formal communications channel to facilitate real-time information sharing regarding election integrity.

C. Candidate Verification

Twitter serves the public conversation by promoting health and earning the trust of the people who use our service. We cannot succeed unless the American people have confidence in the integrity of the information found on the service, especially with respect to information relevant to elections and the democratic process. To promote transparency and assist our stakeholders in identifying messages from elected officials and those who are running for office, we have made a concerted effort to verify all major party candidates for both federal and key state positions. Through verification — a blue badge that appears next to a person’s Twitter handle throughout the service — we let people know that accounts of public interest are the authentic accounts.

D. Election Labels

In addition, we developed a new U.S. election label to identify political candidates. The label includes information about the office the candidate is running for, the state the office is located in, and the district number, if applicable. Accounts of candidates who qualified for the general election and who ran for governor or for the U.S. Senate or House of Representatives displayed an icon of a government building. These new features are designed to instill confidence that the content people are viewing is reliable and accurately reflects candidates’ and elected officials’ positions and opinions.

In total, Twitter identified 1,025 accounts as candidates for the U.S House of Representatives, the U.S. Senate, and state governors. Of the identified accounts, Twitter issued labels to 95% of the identified candidate accounts, or 976 labels. One percent of identified candidates rejected the election label (12 in total) and 4% did not qualify for the label. In the week leading to the elections, individuals on Twitter viewed approximately 100 million impressions of labeled accounts each day, with nearly 13% of U.S. election conversations on the service containing a candidate with a label.



E. Civil Society

Consistent with our longstanding commitment to serving the public conversation, we partnered with experts at the University of Oxford and Leiden University to better evaluate our work on conversation health, focusing on informational echo chambers and unhealthy discourse on Twitter. This collaboration will also enable us to study how exposure to a variety of perspectives and opinions serves to reduce overall prejudice and discrimination. While looking at political discussions, these projects do not focus on any particular ideological group and the outcomes will be published in full in due course for further discussion.

Partnerships with civil society organizations are invaluable in developing our understanding of the behavior of bad-faith actors and responding quickly to potential threats. As we continue to award grants from the \$1.9 million of advertising revenue received from Russia Today and Sputnik to organizations working on elections and civic engagement, we have been able to support capacity building in civil society in furtherance of their election integrity efforts.

Throughout the campaign and on Election Day, our teams were engaged in real-time discussions with civil society organizations who shared their insights on potential issues, enabling us to quickly respond to any activity requiring action.



V. ACTIVITY ON THE SERVICE

The 2018 U.S. midterm elections were the most Tweeted-about midterm elections in history. Twitter facilitated a robust, global conversation regarding the U.S. midterm elections, in which more than 99 million Tweets occurred from the first primaries in March through Election Day. Of the 99 million Tweets regarding the election, 65 million originated in the United States. Of the U.S.-based Tweets, 53% occurred between October 1 through Election Day.

A. Voter Registration Efforts

In addition to the strong discussion we hosted on Twitter, we also collaborated with a number of non-governmental organizations to promote voter registration, civic engagement, and media literacy, including RockTheVote, Democracy Works, TurboVote Challenge, HeadCount, DoSomething, and Ballotpedia.

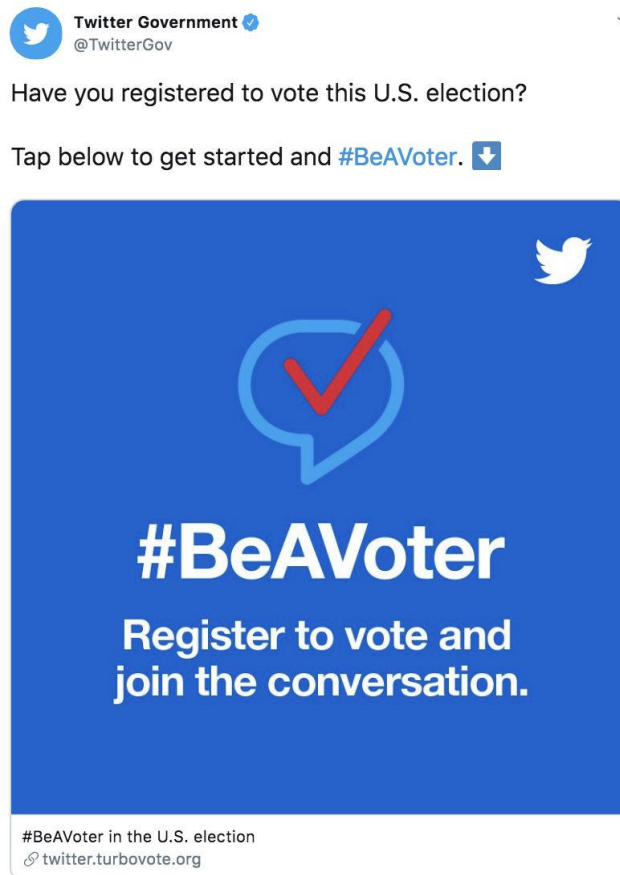
In May 2018, we participated in the TurboVote Challenge Summit along with other industry peers and election nonprofits and presented to over 100 attendees. In July and August, Twitter participated in the leadership committees of TurboVote and National Voter Registration Day for voter engagement efforts.

We deployed three #BeAVoter voter assistance prompts displayed in the home timeline of individuals on the service located in the United States aged 18 and older. Each prompt linked to a nonpartisan, nonprofit managed site for voter assistance that facilitated voter registration and identification of polling locations. The secondary link drove individuals to a Tweet compose window to share with their followers how to register to vote.

Twitter further bolstered the visibility of National Voter Registration Day, a nonpartisan day of action supported by many corporate and nonprofit partners. We saw an increase in Tweets with the primary event hashtag for #NationalVoterRegistrationDay, with a two-fold increase over the number of similar Tweets in 2018, a presidential election year. Of those who Tweeted about National Voter Registration Day, 37.9% had not previously Tweeted about the midterm elections in the six months prior.



We also developed Twitter Emoji to drive meaningful, healthy conversation around civic participation.



Twitter also engaged social influencers to mount a video campaign across @TwitterMovies, @TwitterMusic, @TwitterTV, @TwitterSports, @TwitterWomen. Having a range of social influences for our #BeAVoter video series enabled us to reach a wide variety of people based on interest group. The Tweet from Ariana Grande “thank u, vote” is @Twitter’s second most engaged ever, demonstrating a broad interest in consumer political content.



!

Twitter @Twitter
thank u, vote

5:36 PM · Nov 5, 2018 · Twitter for iPhone

20.7K Retweets 148.5K Likes



B. News Broadcasts

Twitter's News team onboarded over 60 local television stations to Media Studio Producer. Media Studio is a service to manage, measure and monetize broadcast videos on Twitter, in formats such as images, GIFs, videos, and livestreams. Twitter's Media Studio Producer is a new feature that allows publishers to launch professionally-produced live broadcasts on Twitter and Periscope.

Once onboarded, news partners broadcast 32 live debates, forums, and town halls on Twitter relating to the 2018 U.S. midterm elections. At least five of the broadcast debates incorporated questions from the Twitter audience. On Election Day, our partnerships with media outlets resulted in 77 broadcasts on Twitter from 23 national news partners and 112 broadcasts on Twitter from 55 local news partners on Election Day.

C. Monitoring and Identification of Violative Tweets

Individuals on Twitter sent over 99 million Tweets relating to the 2018 U.S. midterm, and Twitter took numerous proactive measures to monitor the service and remove malicious content posted by bad-faith actors. We also received and responded to reports of potential violations of our policies through a variety of channels, including reports through the Partner Support Portal, government entities, and individuals on our service.

1. External Identification of Problematic Content

In advance of the election, Twitter improved the ability of individuals to report problematic content in greater detail at the Tweet or account level. Specifically, Twitter engineered an enhanced spam reporting tool for all individuals using the service to report problematic content. Previously, our spam reporting tool had not provided individuals with specific options for reporting the different types of spam behavior against which Twitter enforces. The additional information provided by individuals on the service enables us to more easily make enforcement actions at scale.



In addition to improving the spam reporting tool, we updated the Partner Support Portal, a tool for allowing partners to rapidly report suspected violations of the Twitter Rules. This improvement facilitates easier reporting from outside partners and further promotes information sharing by tapping into the experience and expertise of active stakeholders. Our goal is to expedite our response to reports from people and organizations active in the election arena. This includes election support organizations, U.S.-based research organizations, and universities and academics who study the spread of misinformation in the media.

Prior to Election Day, we onboarded to the Partner Support Portal more than 10 partners, including the Republican National Committee, the Democratic National Committee, the National Association of Secretaries of State, National Association of State Election Directors, and others. We received 43 reports from our partners, resulting in the removal of thousands of accounts and Tweets in violation of our rules.

2. Internal Identification of Problematic Content

Additionally, Twitter took a number of steps to monitor our service internally. Twitter engaged in proactive monitoring, detection, and surfacing of anomalous behavior related to the election. This monitoring provided visibility into metrics such as Tweet volumes, hashtag tracking, and anomalous behavior at the individual account level. Our monitoring resulted in more than 200 account suspensions, and the removal of more than 5,500 Tweets in violation of our rules. The majority of this content was removed for attempting to suppress voter turnout by sharing misleading or false information about where or how to vote.

Following Election Day, we provided additional monitoring for accounts and hashtags related to races in which a winner had not officially been determined.

3. Examples of Violative Content

The vast majority of violative content we removed from our service on Election Day was voter suppressive content. In general, Twitter looked for behavior that attempted to influence an election by deterring groups of eligible voters, particularly through voter intimidation or providing false information about voting or registering to vote. We removed nearly 6,000 Tweets we identified as attempted voter suppression. Representative examples are below.



Example 1

Username
@handle

BLUE WAVE SMS

text "BLUEWAVE"
to 55463 on election
day and VOTE
from the COMFORT
of your home.*

- VOTE EASY
- VOTE FAST
- VOTE TODAY

 **PROGRESSIVE
TURNOUT PROJECT**
*Service only available for Democrats

AS FEATURED IN

The Boston Globe CNN abc NBC FOX NEWS CBS NEWS The Miami Herald

🗨️ ↻️ ❤️ ✉️

Example 2

Username
@handle

Republicans vote on November 7 in MI-11. Please spread the word.

Lena Epstein  @LenaEpstein
🏠 US House candidate, MI-11

Thank you @VP Mike Pence for stopping by MI-11 today to rally support for our campaign and to encourage voters to go to the polls on November 6th!



🗨️ ↻️ ❤️ ✉️



D. Malicious Accounts Located in Russia

Twitter has seen additional activity on the service affiliated with the Russian Internet Research Agency. As we reported in January 2018, we continue to identify accounts that we believe may be linked to the Russian Internet Research Agency (“IRA”). In September 2018, Twitter’s Chief Executive Officer Jack Dorsey testified on recent activities affiliated with the Russian Internet Research Agency, disclosing that Twitter had suspended a total of 3,843 accounts we believe were linked to the Russian IRA. This number has been updated to 3,613 accounts linked to the Russian IRA.

We continue to build on our contextual understanding of these accounts to improve our ability to find and suspend this activity as quickly as possible in the future, particularly as groups such as the IRA evolve their practices in response to suspension efforts across the industry.

Our ongoing efforts have uncovered an additional 417 accounts that appeared to originate in Russia and are broadly similar in behavior to prior accounts tied to the IRA. Despite these indicators, we cannot render definitive attribution to the IRA for these accounts. While our investigation into the origin of these accounts spanned a period of several months, the vast majority of these accounts were proactively removed by us prior to the 2018 U.S. midterm elections.

On average, these 418 accounts were followed by 385 accounts. These accounts posted in total approximately 929,000 Tweets, of which approximately 81% were in English. In total, 73,398 of the Tweets (approximately 7%) were related to the 2018 U.S. midterms. The most common hashtags were #MAGA (included in approximately 38,000 Tweets), #ReleasetheMemo (included in approximately 38,000 Tweets) and #IslamistTheProblem (included in approximately 18,000 Tweets). In 2015, three of the accounts in this set ran a total of \$5,250 in ads, none of which related to political content; all the ads run by these accounts appeared to have been commercially-motivated spam.

E. Malicious Accounts Located in Iran

In August 2018, we were notified by an industry peer about possible malicious activity on their service originating in Iran. After receiving information from them, we began an investigation on our service to build out our understanding of these networks. We immediately notified law enforcement on this matter as soon as we discovered malicious activity.



We initially identified accounts based on shared indicators such as phone numbers and email addresses. Some of these accounts appeared to pretend to be U.S. persons or news outlets, and discussed U.S. social and political issues. In most cases, the accounts that appeared to suggest a U.S. affiliation were created after the 2016 election. These accounts were in violation of our service manipulation policies, and were engaged in coordinated activity intended to propagate messages artificially across accounts.

Based on our investigation, we believe that these accounts were likely located or created in Iran. This is indicated by, for example, accounts related by an Iranian mobile carrier or phone number or Iranian email address on the account. Although Twitter is blocked in Iran, we may see people active on our service through the use of technologies such as virtual private networks (VPNs).

In September 2018, Twitter disclosed that we suspended 770 accounts for violating Twitter policies. Fewer than 100 of the 770 suspended accounts claimed to be located in the U.S. and many of these were sharing divisive social commentary. On average, these 100 accounts Tweeted 867 times, were followed by 1,268 accounts, and were less than a year old. One advertiser ran \$30 in ads in 2017. Those ads did not target the U.S. and the billing address was located outside of Iran. We remain engaged with law enforcement on this issue.

Since September, through our efforts, we identified and suspended 2,617 additional malicious accounts that we believe had origins in Iran. These accounts posted 4.6 million Tweets, approximately 68% of which were in languages other than English. Of those accounts that tweeted in English, the most popular hashtags were #pakonlinenews (included in approximately 24,000 Tweets), #Palestine (included in approximately 18,000 Tweets), and #Israel (included in approximately 14,000 Tweets). The majority of the content shared by these accounts did not relate to the 2018 U.S. midterm elections: in total, the 2,617 accounts located in Iran tweeted approximately 24,000 times about the 2018 U.S. midterm elections, representing less than 1% of all Tweets from these accounts.

Twitter has been in close contact with our industry peers on this matter and shared detailed information with them about the malicious accounts we believe are located within Iran. This multilateral process of information-sharing will continue, enabling us and our industry peers to work together to better understand and identify malicious activity.



F. Malicious Accounts Located in Venezuela

Twitter recently removed 764 accounts engaged in malicious automation located in Venezuela, some of which Tweeted about the 2018 U.S. midterm election. These accounts Tweeted a total of approximately 985,000 times, and had an average of 591 followers each. All accounts were suspended prior to Election Day, and we notified law enforcement of this activity. Twitter is unable to tie the accounts located in Venezuela to information operations of a foreign government, however, these accounts are another example of a foreign campaign of spammy content focused on political themes, and the behavior we uncovered is similar to that utilized by prior Russian IRA accounts.

Although the vast majority of accounts were removed by November 2017, 176 of the more recently created accounts in this group Tweeted a total of approximately 50,000 times regarding the 2018 U.S. midterm elections. There were no promoted Tweets, and no money spent by these accounts on advertising.



VI. CONCLUSION

The core mission of Twitter is to serve the public conversation. It is why we exist. We must promote and maintain the health of that conversation. The people who use our service must have confidence in the integrity of the information found on the service, especially with respect to information relevant to elections and the democratic process. We continue our efforts to address those threats posed by hostile foreign governments and foster an environment conducive to healthy, meaningful conversations on our service. This work is essential. We will continue to update the public through @Policy, @TwitterSafety, and our dedicated elections hub.